图书情报工作
第64卷 第24期 2020年12月

灰色预测法在情报研究中的应用研究综述*

■ 张晶晶[1,2] 李泽霞[1,2]
[1] 中国科学院文献情报中心 北京 100190 [2] 中国科学院大学经济与管理学院图书情报与档案管理系 北京 100190

摘 要：[目的/意义]灰色预测法可有效处理情报研究中广泛存在的小样本数据，通过对灰色预测法在情报研究中的应用情况进行梳理，总结其在应用过程中存在的不足，为灰色预测法在情报研究中的进一步应用提供参考。[方法/过程]通过综述情报研究中涉及灰色预测法的相关文献，从数据选取、模型构建和解决的问题等方面对情报研究中灰色预测法的应用进行概述，总结当前情报研究中灰色预测法的应用所存在的问题，并提出改进建议。[结果/结论]在方法应用上，已有研究主要采用数列灰预测，且模型集中在单变量灰色预测模型，根据预测对象不同，灰色预测法已经在包括期刊分析、图书馆运行管理、热点主题分析及科研机构评价方面得到了很好的应用，未来可根据预测对象特点及研究目标尝试不同的灰色预测方法，扩宽灰色预测法在其他方面的情报研究问题中的应用。

关键词：灰色预测法 情报研究 数据选取 模型构建 情报研究问题
分类号：G255.51
DOI：10.13266/j.issn.0252-3116.2020.24.012

1 引言

预测作为学术研究中的一种常用手段，可对未来事物的发展情况及发展水平进行科学地推测，进而给当前的研究和工作提供参考，并指导未来的工作规划。在情报领域，研究不仅仅停留于对事物现状的分析，更多时候需要对事物的未来发展趋势进行预测分析，从而为科学决策提供客观量化的依据。以图书馆管理为例，决策者需要根据图书借阅现状及未来可能的图书借阅情况对日后的图书采购进行规划；学校管理者需要根据学科水平现状及未来发展态势对学科发展方向进行调整；科研工作者需要明晰当前研究主题现状及主题发展趋势以不断完善研究内容。

而预测科学经过长时间的发展，积累的预测方法已有300多种，根据预测方式的性质可划分为定性预测和定量预测[1]。其中，定量预测是根据量化数据，结合统计方法和数学模型对事物的未来发展进行预测，主要类别包括时间序列法、因果分析法、灰色预测法、人工神经网络、马尔可夫法等[2]。相较于其他定量预测方法需要建立在大量数据计算的基础上，灰色预测法可有效处理小样本不完全数据，仅需少量数据便可实现较为准确的预测，因而在各类研究中备受关注。

在情报领域，由于数据获取困难等原因，往往会出现数据过少或数据不完整问题，这在一定程度上阻碍了后续的研究分析工作。以大科学装置运行分析为例，大部分大科学装置的运行数据公开程度较低，数据更新不及时，可收集到的运行指标数据通常只有5-6年[3]，远远低于其他预测方法对数据量的要求，难以对其未来的运行情况进行研究。而灰色预测法最少只需4个数据便可实现对小样本数据的有效预测，为情报研究中的小样本数据分析提供了新的方向。事实上，早在1991年就已经有学者在情报研究中应用灰色预测法对科技发展趋势进行预测[4]。近年来，灰色预测法不断推陈出新，但在情报研究中的应用却稍显不足。绝大多数研究只是对基础的灰色预测模型进行简单套用，应用方法单一，缺乏对这一预测方法的系统研究，且预测对象较为局限，用以解决的情报研究问题有待进一步拓展。

为更好地梳理灰色预测法在情报研究中的应用，本文以Web of Science数据库和中国知网为数据来源，

* 本文系中国科学院文献情报中心成立70周年项目（项目编号：院1111）和2020年中国科学院文献情报能力建设专项“科技领域战略情报研究咨询体系建设”（项目编号：E0290001）研究成果之一，受“中国科学院文献情报中心成立七十周年主题论坛与纪念文集出版”项目资助出版。
作者简介：张晶晶（ORCID：0000-0003-2911-2797），硕士研究生；李泽霞（ORCID：0000-0001-7749-4094），研究员，博士，硕士生导师，通讯作者，E-mail：lizexia@mail.las.ac.cn。
收稿日期：2020-11-05 修回日期：2020-11-26 本文起止页码：106-114 本文责任编辑：王传清

分别限定研究方向为"Information Science & Library Science"，文献类别为"图书情报与数字图书馆"，以 SU = (Information Science & Library Science)) AND (TS = ((grey predict *) or (grey forecast *) or (gray predict *) or (gray forecast *) or "GM(1,1)" or "GM(1, n)")) 和 SU = '灰色预测' + '灰预测' + 'GM' 为检索式分别进行检索(检索时间为 2020 年 11 月 15 日)，通过人工阅读剔除与情报研究明显无关或不涉及灰色预测应用的文献，同时对参考文献中出现的有价值文献进行回溯检索，最终整理得到文献 44 篇。本文以这些文献为基础，对情报研究中应用的灰色预测方法及应用这一方法解决的问题进行梳理，并对实际应用中存在的问题进行分析并提出建议，进而为今后情报研究中灰色预测法的应用提供参考。

2 灰色预测法概述

1982 年，灰色系统理论被提出[5]，这是一种解决信息不完备系统的数学方法，可用于分析小样本、不完全信息的系统问题。灰色预测法则是灰色系统理论中的一种方法，是对仅有部分已知信息的灰色系统进行预测，其应用过程主要涉及两个步骤，分别是数据选取以及模型构建。根据数据选取的方式可将灰色预测法划分为多种类型，不同类型的灰色预测法适用于不同特征的数据序列。而模型构建则是通过已经选取的数据建立相应微分方程，从而根据建立的灰色预测模型得出结果，可根据所涉及的变量数进行简单划分。

2.1 数据选取

灰色预测法的第一个主要步骤是对预测对象的数据进行选取处理。根据邓聚龙教授在《灰理论基础》[6]中对灰色预测法的分类，不同类型的灰色预测法的区别在于其构建灰色预测模型所需的数据序列类型有所区别。因此，在灰色预测研究中需要根据预测目标及预测对象特点确定合适的灰色预测法类型以指导预测过程中数据的选取处理工作。目前较为常用的灰色预测法类型包括数列灰预测、灾变灰预测、拓扑灰预测、包络灰预测等[7]，其中数列灰预测是目前最常用的灰色预测法，这一方法直接用观察到的数据进行研究，无需进行数据上的处理，可用于大部分序列。但在异常值的预测中，灾变灰预测表现更佳，这一类型的灰色预测法需要从原始数据中抽取异常值序列进行研究。拓扑灰预测和包络灰预测则更适用于大幅摆动序列，前者是按点集拓扑基选取时点数列，而后者需要构建原始序列的上下边界。

2.2 模型构建

灰色预测法的第二个主要步骤是构建灰色预测模型以对选取处理后的数据进行计算。灰色预测模型发展至今，已积累了许多成熟的模型，但整体而言灰色预测模型可简单分为两类[8]：一类是以 GM (1,1) 模型为代表的单变量灰色预测模型；另一类是以 GM (1,N) 为代表的多变量灰色预测模型，根据研究涉及的变量数可选取不同类型的灰色预测模型构建预测。

2.2.1 单变量灰色预测模型

单变量灰色预测模型是通过单变量的一阶微分方程模型揭示其内在发展规律，主要应用于单一时间序列的建模与预测[9]。邓聚龙提出的 GM (1,1) 模型[5]是最为经典的单变量灰色预测模型，该模型兼具了指数、差分、微分等数学性质。但已有研究表明，GM (1, 1)模型在部分预测条件下会出现较大偏差[10]，因此后续衍生出了一系列改进的 GM (1,1)模型，如刘思峰将实用化缓冲算子引入 GM (1,1)模型[11]，谢乃明等[12]提出了 DGM (1,1)模型，并在 DGM (1,1)模型的基础上更进一步提出了能够解决非齐次近似指数序列预测问题的 NDGM (1,1)模型[13]。

2.2.2 多变量灰色预测模型

多变量灰色预测模型有效克服了单变量灰色预测模型只考虑一个特征参数或者单独考虑几个特征参数发展变化的问题，但是目前多变量灰色模型还主要用于分析系统特征序列与相关因子序列之间的紧密性[14]，较少地应用于预测中。GM (1,N)模型[15]考虑了系统是由多个相互影响的因素组成的情况，是经典的多变量灰色预测模型。但 GM (1,N)模型求解相对复杂，为解决这一问题并提高预测精度，产生了诸如 MGM (1,N)模型[16]、灰色卷积 GMC (1,N)模型[17]、KGM (1, N)模型[18]等多变量灰色预测模型。

3 情报研究中应用的灰色预测法

3.1 数据选取

从数据选取角度分析，目前在情报研究中基本是采用数列灰预测，即直接用观察到的序列进行预测而无需做额外的数据选取和处理。如 X. Dong 等[19]基于收集到的 2004 - 2013 年 TOP 20 的管理学期刊的平均影响因子预测了 2014 - 2018 年平均影响因子，葛凡[20]将采集到的 2007 - 2016 年的图书借阅量直接用于 GM(1,1)模型的构建和预测，苏光耀[21]则应用数列灰预测基于 10 年的论文数量、总被引频次及篇均被引频次进行预测。

已有部分研究注意到数列灰预测在处理非单调序列上表现不佳,尝试采用其他灰色预测类型进行预测,实现对数据的抽取处理。张亮等[22]指出图书馆购书量是典型的波动序列,在研究中对某小型图书馆购书量分别构建了上下边界,应用包络灰预测进行分析;张亮等[23]在原有二维包络预测法的基础上引入时间变量,提出三维包络灰预测实现对图书馆文献知识元数据的预测,这一方法在原有构建上下边界的基础上考虑时间的变化,避免了包络范围可能出现重叠的情况,提高了对波动数据预测的精度。

相比于数列灰预测,诸如灾变灰预测、拓扑灰预测等其他常用灰色预测类型需要对数据进行选取和处理,推广应用程度较低,因而尚未在情报研究中有所应用。

3.2 模型构建

从已有研究中灰色预测模型的应用来看,构建的模型类型涵盖了单变量灰色预测模型和多变量灰色预测模型。

3.2.1 单变量灰色预测模型

单变量灰色预测模型是灰色预测法中使用最为广泛的模型,在情报研究中也是如此,已有研究中所使用的灰色预测模型基本集中于单变量灰色预测模型,特别是GM(1,1)模型。

如朱琳[24]对图书馆绩效的预测、张宁[25]对数据研究热点主题的预测、宋妍[26]对图书馆文献借阅量的预测等均是应用GM(1,1)模型进行预测。GM(1,1)模型虽然应用简单,但对预测数据的限制较多,只适合对指数型数据进行拟合预测,已有研究却往往忽视这一限制,在未分析数据是否具有指数特性的情况下就简单套用这一模型进行分析。而且GM(1,1)模型虽然短期预测效果较好,但在长期预测中,可能会出现预测值远大于实际值的情况,并不适合直接应用于长期预测中。

针对GM (1, 1)模型长期预测效果不佳的问题,部分学者在研究中引入新陈代谢GM(1,1)模型进行预测研究。新陈代谢GM (1, 1)模型是在GM (1, 1)模型的基础上考虑了时间产生的随机扰动因素,将预测到的下一年数据置入原始序列末位,并去掉最老的数据,形成新的时间序列,从而在一定程度上提高了对序列长期预测的精度。钱玲飞和杨建林[27]在学科创新力的预测分析中证实了新陈代谢GM(1,1)模型在长期预测中的精度要明显高于直接应用GM(1,1)模型。Y. Chen[28]同样采用新陈代谢GM(1,1)模型处理预测过程中时间所产生的扰动,实现对图书馆数据服务论文数量的有效预测。此外,基于新陈代谢GM(1,1)模型的预测研究还有陈英等[29]对图书借阅数据的预测,谭博等[30]对全民阅读调查相关数据的预测。常亮等[31]在科技产出的预测研究中应用的等维灰数递补GM(1,1)同样也属于新陈代谢模型。但值得注意的是新陈代谢模型本质仍为GM(1,1)模型,属于指数模型,仅适用于对指数型序列的预测。

鉴于GM(1,1)模型对于非平滑数据预测能力较弱,通过对原始数据进行平滑处理可有效提高GM(1,1)模型的预测精度。徐扬等[32]在情报学热点主题的预测中就在预测前先用对数算子预处理数据,再将处理后的数据代入GM(1,1)模型。将数据对数化是一种较为简单的预处理方法,操作简便,但光滑效果并不明显。施亮[33]则是通过构建4种弱化缓冲算子作用于初始数据之上,弱化了冲击扰动带来的初始数据失真,有效地提高了对高校潜力学科的预测精度。弱化缓冲算子作为目前预测中常见的数据处理方法,在数据光滑处理上表现优秀,通过组合构建不同的缓冲算子可进一步提升数据光滑度。而付新贺等[34]从数据预处理和模型构建两个角度出发,一方面引入差值加权光滑处理,大大提升了数据光滑性,另一方面,针对传统灰色预测模型会引入误差的问题,研究采用欧拉修正的方式进行弥补,改进后的灰色预测模型能够对光滑度较低的图书馆借阅量进行准确预测。

事实上,在预测研究中所涉及的数据多为非齐次数据,仅依靠GM(1,1)模型以及对数据进行光滑处理、误差修正等方式生成的预测结果表现均不如人意。为有效处理非齐次数据,NDGM(1,1)模型被提出,S A. Javed等[35]将这一模型引入情报研究中,运用NDGM(1,1)模型基于美、中、印、巴四国2005-2015年的出版物数量预测2016-2025年的出版物数量,证明了这一方法可很好地处理非齐次数据,对于非齐次序列有着优于GM(1,1)模型的预测精度。

此外,也有研究尝试将灰色预测法与其他预测方法结合起来以进一步提高预测精度。部分研究为提高对非线性数据的处理能力,将GM(1,1)模型与神经网络方法相结合,对图书借阅量[36]、图书副本量[37]进行预测。也有研究将回归方法应用于灰色预测中以进一步提高预测精度,如王居平等[38]分别用二次回归和GM(1,1)模型对图书购置经费进行预测,再应用诱导有序加权平均(Induced Ordered Weighted Average, IOWA)算子将两种预测结果进行组合;刘素兵等[39]将

GM(1,1)和线性回归预测模型相结合,采用"误差平方和最小"作为最优准则实现对图书借阅量的预测。为解决数据波动较大会降低灰色预测精度的问题,马尔科夫链被引入灰色预测中,段玮弘[40]应用GM(1,1)-马尔科夫模型预测了人均借阅册数和馆藏利用率。包红等[41-42]更进一步,将无偏GM(1,1)与马尔科夫链预测方法相结合,分别对图书丢失率和图书借阅量进行预测。从预测结果上看,组合后的预测精度均有一定程度的提高,拓宽了预测方法的适用范围,但组合的模型还只是停留在GM(1,1)模型上。尽管单变量灰色预测模型在情报研究中的应用已相对丰富,但单变量灰色预测模型本身无法考虑与预测变量的相关因素,在一定程度上影响了预测的科学性和准确性。

3.2.2 多变量灰色预测模型

多变量灰色预测模型因求解困难,还较少地应用于各类预测研究中[43],但近年来,也有少量研究尝试应用多变量灰色预测模型进行预测,综合考虑了影响预测对象的有关因素,提高了预测精度。

涉及多变量灰色预测模型的研究有张文德等[44]对人民调解信息的分析,该研究应用GM(1,N)模型对全国化解纠纷件数进行预测,同时为了提高预测精度,研究还尝试将GM(1,N)模型与回归方法结合,实现更为有效的预测分析。此外,滕婕等[45]则将新陈代谢思想与GM(1,N)模型结合,并引入马尔科夫链预测方法,构建了GM(1,N)-马尔科夫动态模型,针对"长生生物疫苗事件"和"北京红黄蓝幼儿园虐童事件"两条热门负面舆情话题,分析了百度指数的未来趋势并进行验证,结果表明这一方法相较于GM(1,N)模型有更好的适用性。

但整体而言,在情报研究中多变量灰色预测模型应用还是相对较少,且应用的方法本质还是将GM(1,N)模型与其他预测方法的简单组合。

4 灰色预测法解决的情报研究问题

从已有研究中灰色预测法的应用来看,灰色预测法已经在包括期刊分析、图书馆运行管理、热点主题分析及科研机构评价等方面得到了很好的应用。

4.1 期刊分析

在情报研究中,对于期刊及其特征的预测分析可有效反映期刊的发展趋势,进而为期刊工作的未来决策提供参考和思路。

灰色预测法在情报研究中的早期应用基本集中于对期刊指标的分析。如陈泉[46]基于1990-1999年期刊出版种数,应用灰色预测法对2000-2005年的期刊出版种数进行了预测;姚红[47]根据1997-2000年中国科技期刊影响因子平均值预测了2001-2004年的影响因子平均值;王筠[48]构建了半衰期指标对《数学学报》的老化情况进行度量,并基于5年已知数据应用GM(1,1)模型预测了这一期刊的未来老化情况。

但灰色预测法对于期刊的研究应用主要集中在2010年以前,预测对象基本为期刊数、影响因子以及老化参数等传统期刊指标。近年来,期刊量化指标不断趋于丰富,却未有见到灰色预测法在这些新指标上的应用。

4.2 图书馆运行管理

图书馆的运行管理一直是学者关注的焦点问题,研究不只局限于对图书馆运行过程的量化指标进行现状分析,越来越多的学者通过对相关指标数据进行预测以指导未来的图书馆规划。灰色预测法是图书馆运行管理研究中最为普遍的方法之一。

常见的研究内容包括对图书借阅量的预测[49-52],这类研究通常是根据前6-8期的图书借阅数据,应用GM(1,1)模型实现灰色预测。此外,还有研究应用灰色预测法对图书馆经费支出[53-54]、图书采购量[22,55]、图书馆运行绩效[24]进行预测。近年来,随着电子图书的普及,也有学者应用灰色预测法对图书馆电子图书的点击量、下载量进行研究[56]。此外,还有包红等[41]应用灰色预测法预测了九江学院图书馆的新旧图书丢失率,刘驰[37]预测了图书的拒借率和平均复本量等。

尽管对于图书馆的灰色预测应用出现时间早,发展时间长,但预测的对象还基本停留在对图书借阅量、图书采购量、图书馆经费等传统运行管理指标的简单预测上。

4.3 热点主题分析

对某一领域或学科进行热点主题预测可对该领域学科的未来发展情况做出归纳总结,为有关主题的研究方向及思路提供一定的参考。

2010年以前,对于热点主题的灰色预测研究更多集中于对有关文献数量的研究[57-58],学者以数据库检索到的相关主题论文数来表征该主题的研究热度,这类研究通常仅是对某一具体学科或领域的整体性分析。

随着自然语言处理技术的成熟,大量主题识别及处理方法被引入情报研究中,对于热点主题的灰色预测研究也从早期的基于文献量转向基于关键词、主题词的预测分析。徐扬等[32]应用GM(1,1)模型对情报

学领域的 9 种主要期刊在 2010 – 2014 年发表论文中最具代表性的 8 个关键词的数据进行计算，预测了这些关键词所代表的热点主题在未来两年可能出现的频率。张宁等[25]聚焦数据研究主题，运用灰色预测法对 2012 – 2016 年 18 种图书情报学期刊刊载的相关论文中的 6 个热点关键词簇进行计算，预测这些关键词簇对应的研究主题在未来两年的研究热度，进而在预测的基础上对数据领域的未来热点进行探讨。王雨等[59]以主题词频数为预测对象，采用 GM(1,1)模型对 CNKI 数据库中 2010 – 2018 年智慧图书馆领域的四大主题的词频数进行定量预测，实现了对四大主题的趋势分析。

值得指出的是，热点主题研究是近年来灰色预测法的应用重点，然而新兴主题分析同为学科主题研究的分支，对这一角度的灰色预测分析还尚未见到。

4.4 科研机构评价

随着国家对建设一流大学和一流学科的不断重视，国内有关机构纷纷开展科研机构及其学科建设的相关评价工作，以指导未来的科研机构发展及学科布局，同时也有大量学者为服务于有关决策者的规划需求，提高科研机构资源配置效率，对机构建设及部分学科的发展态势进行预测评价。

在科研建设方面，董奋义等[60]采用 GM(1,1)模型对高校专利申请受理数、高校科技活动人员、高校 R&D 机构数、高校科技经费筹集额等 9 项指标进行预测；施亮[33]则是借助灰色预测法，基于《武书连中国大学排行榜》中 2012 – 2018 年排名前 10 的高校科研竞争力指数分析了 2019 – 2021 年的科研竞争力指数和变化趋势。

对潜力学科进行预测也是近年来灰色预测法在情报研究中应用的主要方向之一。秦萍等[61]以南京航空航天大学未进入 ESI 排名的 8 个学科为例，基于 2004 – 2013 年 InCites 数据建立 GM(1,1)模型，预测未来若干年高校潜力学科高水平论文的发展趋势，通过总被引频次预测值与 ESI 数据库中的最低值进行比较，得到高校潜力学科与世界前 1% 科研机构的差距，进而预估潜力学科入围 ESI 世界前 1% 的时间。戴莹[62]同样以 ESI 和 InCites 数据库为数据源，以灰色预测法为基础，对合肥工业大学潜力学科的部分指标进行预测与分析。苏光耀[21]则是对浙江工商大学的潜力学科应用灰色预测法进行预测，从而为高校的学科发展提出建议。

不过，科研机构评价相对复杂，包含了科研、教学、基础设施等多个方面，上述的研究均是围绕科研建设过程，还未将灰色预测法应用于其他角度的科研机构评价中。

5 应用中存在的问题及改进

5.1 数据选取相对单一

目前，应用于情报研究的灰色预测法类型基本属于数列灰预测，这意味着在数据的选取上较为单一，其他类型灰色预测法尚不多见，仅有文献[22-23]涉及了包络灰预测。数列灰预测是根据观察到的序列数据直接建立模型进行预测，虽在多数情况下均可使用，但在异常或者波动幅度较大的数值预测上表现不佳，相比之下，灾变灰预测、拓扑灰预测等类型更具优势。

灾变灰预测是对异常值的时序进行预测，主要用于预测异常值在未来发生的时间，这一方法已在其他领域广泛应用，如农业领域用于预测虫害发生时间、经济领域用于预测价格的异常值出现时间。在情报研究中同样存在大量异常数据，如搜索引擎的关键词搜索量在某些日期出现大幅增加的情况，如果能对这些异常值下一次出现时间节点进行有效预测，可以提前对系统的运行能力进行部署。

同样，不同于数列灰预测主要适用于单调序列，拓扑灰预测主要针对大幅度波动序列。在情报研究中，网络舆情跨域周期波动便与之相关，谢科范等[63]指出网络舆情主要经历潜伏期、萌动期、加速期、成熟期、衰退期 5 个阶段，换而言之，当某一网络舆情具体到实际关键词搜索量，其搜索量便是典型的大幅度波动序列，部分还可能具有周期性。“长生生物疫苗事件”和“北京红黄蓝幼儿园虐童事件”的百度指数就属于典型的波动数据[45]，如若在研究中尝试采用拓扑灰预测可更为准确地获取序列的未来趋势。除此之外，对于图书馆借阅量、特定主题的关键词频数等也是情报研究中常见的波动数据。

此外还有包络灰预测、系统灰预测等其他类型的灰色预测法在部分情境下也有优于数列灰预测的表现，实际研究中需根据预测对象的特点及预测的目标选取合适的数据序列进行处理建模。

5.2 模型构建相对基础

当前在情报研究中的灰色预测更多关注单变量灰色预测模型，特别是 GM(1,1)模型。尽管已有部分研究涉及了灰色预测模型的改进，也基本是停留于对 GM(1,1)模型与其他数据处理或量化预测方法的简单组合，其他类型的灰色预测模型并不多见，模型构建相对

基础,这在一定程度上影响了预测结果的精确度。虽然已有部分研究关注到模型所适合的数据类型,但大部分也仅停留在对于数据的预处理,并没有针对数据特点选取其他更为适合的灰色预测模型。相比于 GM(1,1)模型主要适用于指数型数据,DGM(1,1)模型、NDGM(1,1)模型等这类对数据指数特征要求较低的模型都是未来研究中可进一步考虑的灰色预测模型。

同时,邓聚龙教授[64]指出在预测研究中大多数情况是非线性数据,诸如文献的增长、被引频次变化在更多情况下都是非线性序列。然而 GM(1,1)模型处理非线性问题能力较弱,尽管已有部分研究尝试将 GM(1,1)模型与神经网络方法结合以有效处理非线性数据,并取得了相对令人满意的结果,但如果能在未来的研究中引入适合处理非线性数据能力的灰色预测模型,并将其与神经网络方法相结合,可进一步提高对非线性数据的预测能力,常见的非线性灰色预测模型有非线性 GM(1,1,α) 模型、非线性优化 GM(1,N)模型、KGM(1,N)模型等。

多变量灰色模型的应用同样值得关注。在实际预测中,预测对象往往与众多因素相互关联,单变量灰色预测模型仅考虑预测对象这一因素,无法反映出各个变量之间的相互影响及协同发展,而多变量灰色预测模型能够探究多个变量之间的关系,因此,在情报研究中可以考虑引入多变量灰色模型进行预测。文献[35]是通过 NDGM(1,1)模型实现对各国出版物数量的预测,事实上,一国的出版物数量与国家科研实力、科研投入、科研工作者数量等因素均可能相关,如若在预测过程中尝试引入多变量灰色模型,将这些可能涉及的因素纳入模型中,在一定程度上可进一步提高预测结果的科学性、准确性。

5.3 用以解决的问题相对局限

目前灰色预测法在情报研究中的应用基本局限于期刊分析、图书馆运行管理、热点主题分析和科研机构评价的部分内容上,用以解决的问题相对有限,研究内容有待进一步丰富尝试。

一方面,对于这 4 个方面的应用均可再进一步深入。目前对于期刊的研究分析基本还停留于最简单的期刊数量、老化参数、影响因子的预测上,近几年随着期刊量化指标的不断丰富,已经衍生出了诸如 SJR 指数(又称期刊声望指数)[65]、Eigenfactor Scores(特征因子)[66]、I3 指标[67]等期刊评价指标,同时 Altmetrics 的出现也给期刊的分析评价带来了一系列替代计量指标,未来对于期刊的预测分析可从这些新的量化指标出发进行分析。对于图书馆运行管理过程的预测研究也基本集中于借阅量、经费支出这些最简单的方面,但是,近年来,伴随着数字化程度的不断提高,公众知识水平的不断提升,图书馆的运行管理所涵盖的内容不断增多,例如活动开展、网络平台构建、文献传递服务等,这些方面均会产生大量过程性或结果性数据,如果能够从这些角度进行灰色预测分析将会为未来图书馆的运行管理工作提供参考。目前,热点主题的研究是灰色预测的应用焦点,却未见这一方法在新兴主题预测上的应用,相较于热点预测,新兴主题预测强调的是新颖性及发展潜力[68],对于调整研究规划,完善科研决策具有重要意义,值得在灰色预测研究中加以尝试。而在科研机构评价中,目前的灰色预测法应用基本是围绕对高校潜力学科的研究,重点关注高校的科研建设,事实上,科研机构评价除了包含对其科研水平及学科发展的评价,还涵盖了对其教学水平、国际化水平、基础设施水平、知识转移能力等方面的评价,可将灰色预测法应用至高校评价中的其他方面,实现对高校科研能力的综合评价。

另一方面,也可拓宽灰色预测法在情报研究中的应用范围。学者影响力预测可为学术决策提供参考,有助于学者科研事业的规划,是近年情报预测研究中关注的角度,其预测方法基本集中在回归分析和神经网络[69],但对于较为年轻的学者,其有关影响力指标数据量相对较少,应用灰色预测法进行预测将是一个不错的选择。灰色预测法在颠覆性技术预测上的应用同样值得尝试,对颠覆性技术进行预测将有助于国家在日益激烈的技术竞争中抢占先机,目前的预测方法以神经网络和时间序列预测方法为主[70],然而近年来,颠覆性技术从技术理论形成到技术广为使用所需的时间不断缩短,这使得颠覆性技术预测中所涉及的指标数据通常只有短短几年,数据量较低,引入灰色预测法对这些小样本指标数据进行预测可有效提高预测准确性。此外,还有跨学科预测、科研合作预测、学术新星预测等都是近年情报预测研究中关注的焦点,但目前对于这些角度的研究均未涉及灰色预测法,在未来的应用中可尝试将灰色预测法应用于这些方面,拓展灰色预测法在情报研究中的应用范围。

6 结论

本文通过对情报研究中应用的灰色预测方法及这一方法解决的情报研究问题进行梳理,分析了目前研究中灰色预测法的应用所存在的不足。目前,灰色预

测法已在多方面的情报研究中得到了大量应用,方法应用上主要集中在数列灰预测,模型则以单变量灰色预测模型及其简单的组合应用为主,在包括期刊分析、图书馆运行管理、热点主题分析和科研机构评价 4 个方面均得到了较为广泛的应用,并取得了不错的效果,在一定程度上解决了数据量偏低所带来的预测难问题,为研究者、决策者提供了大量参考性意见。未来,在灰色预测法的应用中,一方面可在已有研究内容上进一步深入,对期刊、图书馆运行、热点主题、科研机构方面的研究内容进行延伸,应用灰色预测法对其他研究指标进行分析,另一方面也可尝试扩展至新的应用领域,在诸如学者影响力、颠覆性技术、跨学科性、科研合作等方面的预测加以应用。同时,在研究中根据预测对象的特点及研究目标,可尝试引入一系列更为合适的预测类型,如灾变灰预测、拓扑灰预测、包络灰预测等对异常值或波动性序列进行分析,以不同的方式选取和处理数据;可尝试引入新的灰色预测模型,如非线性、非齐次模型,提升预测结果的准确性;进一步深化多变量灰色预测模型在情报研究中的应用,将可能影响预测对象的其他因素纳入分析。相信随着灰色预测法的不断丰富,灰色预测法在情报研究中的应用会进一步深化,预测的结果将会更加精确合理。

参考文献:

[1] 鞠雨宏,张焕伟. 经济预测方法综述与选择[C]//吉林省财政科学研究所. 财金观察. 长春:吉林省财政科学研究所, 2020: 159-168.
[2] 袁军鹏. 一种两阶段数量组合预测方法及实证[J]. 情报学报, 2014, 33(9): 910-915.
[3] 王婷, 陈凯华, 卢涛, 等. 重大科技基础设施综合效益评估体系构建研究——兼论在 FAST 评估中的应用[J]. 管理世界, 2020, 36(6):218-241,260.
[4] 孟骏, 孟庆湖. 利用灰色模型预测科技发展趋势[J]. 情报理论与实践, 1991, 14(6): 13-15.
[5] DENG J L. Control problems of grey systems[J]. Systems & control letters, 1982, 1(5):288-294.
[6] 邓聚龙. 灰理论基础[M]. 武汉: 华中科技大学出版社, 2002.
[7] 陈飞. 灰色预测模型及应用[J]. 商情(教育经济研究), 2008, 2(5): 112.
[8] PAN W W, JIAN L R, LIU T. Grey system theory trends from 1991 to 2018: a bibliometric analysis and visualization[J]. Scientometrics, 2019, 121(3): 1407-1434.
[9] 邓聚龙. 灰色系统[M]. 北京: 国防工业出版社, 1985.
[10] 熊和金,徐华中. 灰色控制[M]. 北京: 国防工业出版社, 2005.
[11] LIU S F. The three axioms of buffer operator and their application to GM(1,1) prediction[J]. Journal of grey system, 1991, 3(1): 39-48.
[12] XIE N M, LIU S F. Research on discrete grey model and its predictive result[J]. Journal of systems engineering, 2006, 21(5): 520-523.
[13] 谢乃明, 刘思峰. 近似非齐次指数序列的离散灰色模型特性研究[J]. 系统工程与电子技术, 2008, 30(5): 863-867.
[14] TONG J, FU J, WANG Q, et al. A grey estimation method for the seismic intensity[J]. Journal of grey system, 2011, 23(3): 251-256.
[15] DENG J L. Introduction to grey system theory[J]. The journal of grey theory, 1989, 1(1): 1-24.
[16] 翟军, 盛建明, 冯英浚. MGM(1,n)灰色模型及应用[J]. 系统工程理论与实践, 1997,17(5): 110-114.
[17] TIEN T. The indirect measurement of tensile strength of material by the grey prediction model GMC(1, n)[J]. Measurement science & technology, 2005, 16(6): 1322-1328.
[18] MA X, LIU Z B. The kernel-based nonlinear multivariate grey model[J]. Applied mathematical modelling, 2018, 56(4): 217-238.
[19] DONG X, ZHANG S Q. Impact factor dynamic forecasting model for management science journals based on grey system theory[EB/OL]. [2020-10-20]. https://file.scirp.org/Html/54756_54756.htm.
[20] 葛凡. 基于灰色系统模型的图书借阅量预测分析——以青岛科技大学图书馆为例[J]. 教育教学论坛, 2018, 10(46): 106-109.
[21] 苏光耀. 基于灰色系统理论的高校潜势学科发展预测研究——以浙江工商大学为例[J]. 情报探索, 2018(5): 8-14.
[22] 张亮, 赵云. 图书馆采购文献的灰色预测模型构建[J]. 情报杂志, 2007, 26(6): 131-133.
[23] 张亮, 杨溢. 论基于三维包络灰预测与知识元理论的图书馆文献采购优化[J]. 图书馆学研究, 2013(8): 45-48.
[24] 朱琳. 灰色理论在高校图书馆部门绩效评价中的应用[J]. 产业与科技论坛, 2014, 13(13): 134-135.
[25] 张宁, 袁勤俭, 沈洪洲. 我国图情领域有关数据研究的态势分析——基于灰色预测算法[J]. 情报杂志, 2017, 36(8): 156-160.
[26] 宋妍, 肖希明. 基于 GM(1,1)模型的大学图书馆文献借阅量预测与分析[J]. 云南师范大学学报(自然科学版), 2018, 38(6): 35-39.
[27] 钱玲飞, 杨建林. 基于 GM(1,1)模型的学科创新力预测[J]. 情报科学, 2012, 30(4): 530-534.
[28] CHEN Y. GM(1,1) Based analysis of national research tendency on hotspot of library's data service in last decade[C]// Proceedings of the 3rd international conference on contemporary education, Social Sciences and Humanities (ICCESSH 2018). Paris: Atlantis Press, 2018: 1265-1268.
[29] 陈英, 王秀山. 基于灰色系统理论的农业院校院系纸质图书借阅管理研究[J]. 科技视界, 2013(3): 114-116.

[30] 谭博，雷润玲，李航. 我国全民阅读调查分析与趋势预测[J]. 图书情报工作，2017，61(6)：87－91.

[31] 常亮，唐永林. 基于灰色系统理论的我国科技活动分析与预测[J]. 现代情报，2006，26(11)：35－39,42.

[32] 徐扬，孟文霞，李广建. 基于灰色预测模型的情报学热点主题发展预测[J]. 情报科学，2016，34(7)：3－6,30.

[33] 施亮. 基于GM(1,1)改进模型的高校科研竞争力研究[J]. 图书馆研究与工作，2019(1)：11－15.

[34] 付新贺，袁永旭. 基于灰色预测算法的图情数据分析技术[J]. 电子科技，2020，33(7)：67－70.

[35] JAVED S A，LIU S F. Predicting the research output/growth of selected countries：application of Even GM (1，1) and NDGM models[J]. Scientometrics，2018，115(1)：395－413.

[36] 张囡，张永梅. 基于灰色神经网络的图书馆图书借阅量预测[J]. 情报探索，2013(3)：133－135.

[37] 刘驰. 灰色神经网络模型在图书采购复本量预测的应用[J]. 科技创新导报，2013，10(19)：200－201.

[38] 王居平，陈华友. 一种新的图书购置经费的组合预测方法[J]. 情报杂志，2007，26(4)：24－25,29.

[39] 刘素兵，刘海明，苗佳晶，等. 图书借阅量的灰色回归组合模型研究[J]. 云南民族大学学报(自然科学版)，2010，19(3)：170－172.

[40] 段玮弘. 基于灰色－马尔柯夫模型的图书借阅行为流量预测研究[J]. 鲁东大学学报(自然科学版)，2011，27(3)：207－212.

[41] 包红，陈小艳. 高校图书馆图书丢失率的无偏灰色GM(1,1)预测模型研究——以九江学院图书馆为例[J]. 情报探索，2011(7)：13－14.

[42] 包红，刘臻. 高校图书流通量的优化灰导数白化值的无偏灰色GM(1,1)模型预测[J]. 农业图书情报学刊，2011，23(2)：65－68.

[43] TIEN T L. A research on the grey prediction model GM(1,n)[J]. Applied mathematics and computation，2012，38(9)：4903－4916.

[44] 张文德，刘静，张婷，等. 组合预测模型在信息预测服务中的应用——以人民调解数据为例[J]. 情报探索，2019(6)：21－26.

[45] 滕婕，夏志杰，罗梦莹. 基于新陈代谢GM(1,N)马尔科夫模型的动态网络舆情危机预测[J]. 情报科学，2020，38(8)：88－94.

[46] 陈泉. 图书期刊出版种数的灰色预测模型[J]. 图书情报工作，2003，47(8)：50－52.

[47] 姚红. 期刊计量指标的灰色预测研究[J]. 图书情报工作，2004，48(10)：5－7,11.

[48] 王筠. 科技文献老化的灰色预测与分析[J]. 情报理论与实践，2003，26(1)：40－42.

[49] 孔超. 灰色预测模型在高校图书馆采访中的应用[J]. 知识管理论坛，2014(4)：35－39.

[50] 孙宝. 图书馆借阅量预测的灰色系统模型[J]. 华北科技学院学报，2008，5(1)：122－124.

[51] 孙宝,王志丽,刘琳. 基于GM的高校图书馆借阅量预测模型研究[J]. 现代情报，2008，28(4)：186－188.

[52] 周建良. 基于GM的图书需求预测模型研究[J]. 鸡西大学学报，2007，7(3)：92－94.

[53] 李冬梅. 数字化图书馆建设经费投入预测模型及其应用[J]. 网络财富，2010 (11)：135,137.

[54] 王廷满，沈思. 利用灰色系统理论预测图书文献经费[J]. 西安科技学院学报，2002，22(3)：367－370.

[55] 开滨. 基于灰色GM(1,1)模型的图书馆采购预测研究[J]. 情报探索，2014(12)：91－93.

[56] 过仕明，吕莉媛. “互联网＋”背景下高校图书馆协同服务机制研究[J]. 情报科学，2019，37(4)：52－58.

[57] 吴淑玲，艾学涛. 灰色模型在图书馆管理研究预测中的应用[J]. 情报杂志，2006，25(6)：130－132.

[58] 徐曼，张爱军. 1995－2005年中国干旱区研究发展态势及学术影响力分析[J]. 冰川冻土，2007，29(4)：659－663.

[59] 王雨，盛东方，王铮. 我国智慧图书馆研究热点分析与趋势预测[J]. 图书馆论坛，2020，40(6)：39－46.

[60] 董奋义，刘斌. 高校科技产出与投入的灰色关联分析和趋势预测[J]. 昆明理工大学学报(理工版)，2010，35(1)：113－118.

[61] 秦萍，李雪琛，梅秀秀. 高校潜力学科发展预测研究[J]. 情报杂志，2015，34(1)：88－91.

[62] 戴莹. 灰色预测在高校学科发展中的应用——以合肥工业大学为例[J]. 情报探索，2018(9)：13－17.

[63] 谢科范，赵湜，陈刚，等. 网络舆情突发事件的生命周期原理及集群决策研究[J]. 武汉理工大学学报(社会科学版)，2010，23(4)：482－486.

[64] DENG J L. Introduction to grey mathematical resource science[M]. Wuhan：Huazhong University of Science & Technology Press，2010.

[65] FARNEY T. SCImago journal and country rank[J]. Charleston advisor，2009，11(1)：51－54.

[66] BERGSTROM C. EIGENFACTOR：Measuring the value and prestige of scholarly journals[J]. College and research libraries news，2007，68(5)：314－316.

[67] LEYDESDORFF L. Alternatives to the journal impact factor：I3 and the top－10% (or top－25%?) of the most-highly cited papers[J]. Scientometrics，2012，92(2)：355－365.

[68] 白敬毅，颜端武，陈琼. 基于主题模型和曲线拟合的新兴主题趋势预测研究[J]. 情报理论与实践，2020，43(7)：130－136,193.

[69] 夏琬钧，任鹏，陈晓红. 学者影响力预测研究综述[J]. 情报理论与实践，2020，43(7)：165－170.

[70] 石慧，潘云涛，苏成. 颠覆性技术及其识别预测方法研究综述[J]. 情报工程，2019，5(3)：33－48.

作者贡献说明：

张晶晶：论文框架设计，论文撰写；

李泽霞：论文框架设计，论文修改。

Review on the Application of Grey Prediction Method in Information Research

Zhang Jingjing[1,2] Li Zexia[1,2]
[1] National Science Library, Chinese Academy of Sciences, Beijing 100190
[2] Department of Library, Information and Archives Management, School of Economic and Management, University of Chinese Academy of Sciences, Beijing 100190

Abstract: [Purpose/significance] Grey prediction method can effectively deal with small sample data widely existing in information research. This paper analyzed the application of grey prediction in information research and summarized the shortcomings in the application process, laying the foundation for further application of grey prediction method in information research. [Method/process] By reviewing the literatures related to grey prediction method in information research, this paper gave an overview of data selection, model construction and solved problems of grey predicting method in information research, summarized the problems existing in the application of grey prediction method in current information research, and put forward suggestions for improvement. [Result/conclusion] In the application of the method, the research mainly adopts the sequence grey prediction, and the model focuses on the univariate grey prediction model. According to different prediction objects, grey prediction method has been well applied in journal analysis, library operation management, hot topic analysis and evaluation of scientific research institutions. In the future, different grey prediction methods can be tried according to the characteristics of the prediction object and the research objectives, and the application of grey prediction method can be expanded to other information research issues.

Keywords: grey prediction method information research data selection model construction information research issue